

Phenome-wide association study of ovarian cancer identifies common comorbidities and reveals shared genetics with complex diseases and biomarkers

Anwar Mulugeta^{1,2&3*}, Amanda L Lumsden^{1&2}, Iqbal Madakkatel^{1&2}, David Stacey^{1&2}, S.Hong Lee^{1,2&4}, Johanna Mäenpää^{5&6}, Martin K. Oehler^{7&8}, Elina Hyppönen^{1&2}

¹Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of South Australia, Adelaide, Australia

²South Australian Health and Medical Research Institute, Adelaide, Australia

³Department of Pharmacology and Clinical Pharmacy, College of Health Science, Addis Ababa University, Addis Ababa, Ethiopia

⁴UniSA Allied Health & Human Performance, University of South Australia, Adelaide, SA, Australia

⁵Faculty of Medicine and Medical Technology, Tampere University, Tampere, Finland

⁶Cancer Centre, Tampere University and University Hospital, Tampere, Finland

⁷Department of Gynaecological Oncology, Royal Adelaide Hospital, Adelaide, SA, Australia

⁸Adelaide Medical School, Robinson Research Institute, University of Adelaide, Adelaide, SA, Australia

Short title: MR-PheWAS of ovarian cancer

Funding: This research was funded by Medical Research Future Fund of Australia (MRF2007431).

Competing interests: The authors declare that they have no competing interests.

***Corresponding author:**

Dr Anwar Mulugeta

Australian Centre for Precision Health, c/o SAHMRI, University of South Australia, Adelaide, Australia, GPO Box 2471, Adelaide SA 5001.

Email: anwarmulugeta.gebremichael@unisa.edu.au

1 **Abstract**

2 **Background:** Ovarian cancer (OC) is commonly diagnosed among older women who have
3 comorbidities. This hypothesis-free phenome-wide association study (PheWAS) aimed to identify
4 comorbidities associated with OC, as well as traits that share a genetic architecture with OC.

5 **Methods:** We used data from 181,203 white British female UK Biobank participants and analysed
6 OC and OC subtype-specific genetic risk scores (OC-GRS) for an association with 889 diseases and
7 43 other traits. We conducted PheWAS and colocalisation analyses for individual variants to identify
8 evidence for shared genetic architecture.

9 **Results:** The OC-GRS was associated with 10 diseases, and the clear cell OC-GRS was associated
10 with five diseases at the FDR threshold ($p = 5.6 \times 10^{-4}$). Mendelian randomisation analysis (MR)
11 provided robust evidence for the association of OC with higher risk of “secondary malignant
12 neoplasm of digestive systems” (OR 1.64, 95% CI 1.33, 2.02), “ascites” (1.48, 95% CI 1.17, 1.86),
13 “chronic airway obstruction” (1.17, 95% CI 1.07, 1.29), and “abnormal findings on examination of
14 the lung” (1.51, 95% CI 1.22, 1.87). Analyses of lung spirometry measures provided further support
15 for compromised respiratory function. PheWAS on individual OC variants identified five genetic
16 variants associated with other diseases, and seven variants associated with biomarkers (all, $p \leq 4.5 \times 10^{-8}$).
17 Colocalisation analysis identified rs4449583 as the shared causal variant for OC and seborrheic
18 keratosis.

19 **Conclusions:** OC is associated with digestive and respiratory comorbidities. Several variants
20 affecting OC risk were associated with other diseases and biomarkers, with this study identifying a
21 novel genetic locus shared between OC and skin conditions.

22

23 **Keywords:** Ovarian cancer, comorbidities, *TERT*, PheWAS, UK Biobank

24

25 **1. Introduction**

26 Ovarian cancer (OC) is a notorious gynaecological disorder, with only 44% of affected women
27 surviving 5 years.¹ Reducing mortality associated with OC remains a challenge as the poor
28 understanding of the origin and biology of the disease limits opportunities to prevention, and the
29 disease is typically detected at an advanced stage when treatment options are limited.² Approximately
30 90% of OC are epithelial OC and arise from epithelial cells, predominantly of the fallopian tube and
31 endometrium.²⁻⁴ Epithelial OC can be further classified into serous, endometrioid, clear cell, and
32 mucinous subtypes, based on their histological characteristics³ and indeed, the pathogenesis, disease
33 progression, and risk factors are known to vary between OC subtypes.^{3,5} Given that OC is commonly
34 diagnosed in women aged 50 years and above,⁶ many patients present with at least one comorbidity,
35 which can influence treatment decisions, impact prognosis and affect overall survival.^{7,8} Hence,
36 exploring the associations between OC and comorbidities, along with understanding potential shared
37 genetic links between OC and other traits, may offer insights into the underlying etiology of the
38 disease. Such insights could lead to facilitating earlier detection and better management of the disease.

39 Phenome-wide association study (PheWAS) is a powerful approach for testing the association of
40 genetic variants or their risk score with a broad range of phenotypes and disease outcomes.⁹ This
41 approach offers the advantage of being hypothesis-free, allowing it to explore numerous associations
42 without constraints of the number of outcomes. Consequently, PheWAS has the potential to unveil
43 novel genetic associations with diseases.⁹ In this study we use data on genetic OC risk variants and
44 889 disease outcomes from 181,203 female UK Biobank participants to identify comorbidities and
45 clinical characteristics associated with the genetic risk of OC. We test for evidence of a direct
46 association between OC and possible comorbidities using Mendelian randomisation (MR), an
47 approach less susceptible to confounding and reverse causation,¹⁰ and extend these analyses to focus
48 on the genetic risk reflecting distinct histological OC subtypes. We also use the PheWAS approach to
49 explore each OC associated variant individually, which can provide evidence for shared genetic
50 contributions, and provide insight into the underlying disease mechanisms.

51

52 **2. Methods**

53 **2.1. Study population**

54 UK Biobank is a prospective cohort study of half a million participants with rich phenotype and
55 genotype data collected across 22 centres in England, Scotland, and Wales. The baseline assessment
56 was conducted between 2006 and 2010 when the participants were aged between 37 to 73 years.¹¹ A
57 wide range of data was collected, including sociodemographic information, lifestyle factors, physical
58 measures, blood and urine samples for biomarker and genetic profiling, with continuing updates
59 through follow-up assessments and linkage with electronic health records and death registrations.¹¹
60 For the primary analyses in this study, we restricted the sample to 181,203 female participants who
61 were of white British ancestry and were not genetically related (Figure S1) and included male
62 participants in sensitivity analyses (n =156,220). The study was conducted under UK Biobank
63 application number 89630. UK Biobank obtained informed consent from all participants to collect and
64 use linked data for future research use, and the UK Biobank study has ethical approval from the
65 National Information Governance Board for Health and Social Care and Northwest Multicenter
66 Research Ethics Committee (11/NW/0382).

67

68 **2.2. Disease outcomes and biomarkers**

69 We used International Classification of Diseases (ICD version 9 and 10) codes from hospital episode
70 statistics and death registrations to identify disease outcomes (Supplementary methods). We mapped
71 ICD-9/10 codes to phecodes, which are composite groupings reflecting clinically relevant disease
72 outcomes.¹² We defined individuals with the phecode-of-interest as cases, and those without the
73 phecode-of-interest, or any other phecodes from the same disease category as controls
74 (Supplementary methods).^{12, 13} We included all phecodes with at least 200 cases in our analyses,¹⁴
75 resulting in 889 phecodes from 18 disease categories (Table S1).

76 We also included 30 serum and four urine biomarkers, which are commonly used for the diagnosis
77 and monitoring of chronic diseases, as outcome measures (Supplementary methods). Additional
78 analyses were conducted against selected physiological characteristics (namely body mass index
79 (BMI), body fat percentage (BFP), basal metabolic rate (BMR), blood pressures, and measures of
80 lung function (spirometry) due to their known associations with comorbidity risks. The spirometry
81 measures included forced expiratory volume in 1-second (FEV1), forced vital capacity (FVC),
82 FEV1/FVC ratio, and peak expiratory flow (PEF).

83 **2.3. Genetic variants and genetic risk score**

84 We selected 35 genetic variants associated with OC from previous genome-wide association studies
85 (GWAS), including the 12 novel variants identified in the most recent GWAS¹⁵ and 23 variants
86 identified in earlier GWASs (minor allele frequency $\geq 1.4\%$, Table S2, Figure S2).¹⁶⁻²³ We
87 constructed the genetic risk score (GRS) using 31 genetic variants after excluding one variant that had
88 no proxy at $r^2 \geq 0.8$ (rs555025179), and three variants (rs12131772, rs4691139, and rs12938171) that
89 were not replicated in the recent GWAS,¹⁵ which we extracted from the imputed genotype UK
90 Biobank data²⁴ (Table S2). The GRS was constructed for each individual by summing the weighted
91 number of OC risk-associated variants, with the weight taken from the recent OC GWAS.¹⁵ We also
92 constructed risk scores for each OC subtype (Supplementary methods).

93 **2.4. Statistical analyses**

94 We conducted phenome-wide analyses using R-package PheWAS²⁵ which allowed us to fit logistic
95 regression (889 disease outcomes) or linear regression (43 biomarkers/physiological characteristics)
96 against the OC GRS with all models adjusted for age, assessment centre, type of genotyping array,
97 birth location, and 40 principal components (Figure S2). We repeated the phenome-wide analyses for
98 each OC subtype GRS to identify subtype-specific disease associations, and further, conducted the
99 analyses in male participants to understand whether the OC – disease associations involve non-sex-
100 specific mechanistic pathways.

101 The phenome-wide association signals that passed the FDR threshold ($p < 5.6 \times 10^{-4}$ for the diseases
102 and $p < 0.02$ for the biomarkers and physiological measures, Supplementary methods) were taken
103 forward to MR analysis to examine evidence for direct association with OC. We used two-sample MR
104 approach where variant – OC association estimates were taken from the recent OC GWAS,¹⁵ and the
105 variant-outcome association estimates were from the UK Biobank. MR analyses were conducted on
106 signals associated with the overall OC GRS or with any subtype-based GRSs. We used inverse
107 variance weighted (IVW) MR²⁶ as the primary method and conducted sensitivity analyses using
108 weighted median,²⁷ weighted mode,²⁸ MR-PRESSO,²⁹ and MR-Egger,³⁰ with each method working on
109 different pleiotropic assumptions (Supplementary methods). We tested for horizontal pleiotropy using
110 MR-Egger intercept, MR-PRESSO outlier and distortion tests, leave-one-out and leave-block-out
111 methods (Table S3, Supplementary methods). MR analyses were repeated excluding the OC cases. To
112 test the MR assumption that the genetic instrument should not be associated with potential
113 confounding factors, we explored the association between the OC GRSs and age, education,
114 Townsend deprivation index, BMI, physical activity, and alcohol consumption (Supplementary
115 methods, Table S4).

116 We repeated the phenome-wide analyses for each OC genetic variant individually, again conducting
117 analyses separately for men. We assessed the novelty of the PheWAS findings by checking the
118 reported associations in the GWAS Catalog. We considered a locus-trait association as potentially
119 novel if neither the lead variant nor any variant in linkage disequilibrium (LD) with the lead variant
120 (with an $r^2 \geq 0.1$) had previously been reported to be associated with the same or related traits. For
121 those potentially novel associations, we conducted a follow-up colocalisation analysis using the
122 HyPrColoc (Hypothesis Prioritization in multi-trait Colocalisation) tool to assess whether there is
123 evidence of shared causal genetic variants between pairs of traits.³¹ We considered a posterior
124 probability of full colocalisation (PPFC) > 0.64 as indicative of shared genetic etiology between these
125 traits (Supplementary methods).³¹

126 Statistical analyses were performed using STATA SE version 17.1, R version 4.2.1 and PLINK2.³²

127 **3. Results**

128 We included up to 181,203 female white British participants in the primary analyses, of whom 1,863
129 (1.3%) had been diagnosed with OC (Table 1). The prevalence of OC was higher among participants
130 who were older, obese, never drinkers of alcohol, and those with poor general health status at the
131 baseline (for all, prevalence $\geq 1.2\%$, and $p \leq 0.03$). The OC risk GRS was normally distributed and
132 associated with OC as expected, with stronger association in the highest compared to the lowest
133 genetic risk group (F-statistics =23.7, $p = 4.4 \times 10^{-16}$, Figure S3). No associations were found between
134 the risk scores and potential confounders after correction for multiple testing (Table S5).

135 **3.1. Phenome-wide-based Mendelian randomisation**

136 We investigated 889 disease outcomes from 18 disease categories in the phenome-wide analyses. We
137 observed significant associations between the OC GRS and 10 diseases from four disease categories
138 (neoplasms, digestive, respiratory and genitourinary diseases) after FDR correction ($p \leq 5.6 \times 10^{-4}$).
139 The top signals were for “cancer of the female genital organs”, “malignant neoplasm of ovary and
140 other uterine adnexa”, and “malignant neoplasm of the ovary”, which confirms the ‘relevance
141 assumption’ of the MR approach¹⁰ (Supplementary methods) for using the GRS as instrument for OC
142 (for all, $p \leq 4.8 \times 10^{-4}$, Figure 1, Panel A). The GRS was also associated with higher risk of other
143 cancers and abnormal growths including “secondary malignant neoplasm of digestive systems”,
144 “lipoma, including lipoma of skin and subcutaneous tissue” ($p \leq 3.9 \times 10^{-4}$), and “ovarian cyst” ($p = 6.0$
145 $\times 10^{-5}$). In addition, we identified associations with respiratory outcomes including “chronic airway
146 obstruction” and “abnormal findings on examination of the lungs”, and with the digestive system
147 outcome, “ascites” ($p \leq 3.2 \times 10^{-4}$).

148 In analyses using GRSs for specific OC subtypes (Figure 1 Panel B and C), the clear cell OC GRS
149 was associated with additional diseases and increased risks such as “benign neoplasm of uterus”, and
150 “uterine leiomyoma”, “disorders of menstruation”, “gram negative septicemia”, and “postoperative
151 infection” (for all, $p \leq 2.3 \times 10^{-4}$). Overall, in addition to overall OC-related outcomes, 12 other
152 disease outcomes were associated with at least one form of the OC GRS. Notably, none of these
153 associations were detected when we repeated the phenome-wide analyses among men (Figure S4).

154 Additionally, our phenome-wide analyses on 43 quantitative traits, which included serum and urine
155 biomarkers and physiological measures, identified an association of OC GRS with 13 traits.
156 Furthermore, we found an additional 11 traits associated with at least one subtype of OC GRS (Figure
157 S5).

158 We conducted two-sample MR analyses on 12 disease outcomes identified in the PheWAS. For each
159 outcome, evidence of causality was identified for at least one type of OC (Figure 2). We found robust
160 evidence supporting the association between a higher genetic liability of OC and an increased risk of
161 “secondary malignant neoplasm of digestive systems” (OR 1.64, 95% CI 1.33, 2.023), “ascites” (1.48,
162 95% CI 1.17, 1.86), “chronic airway obstruction” (1.17, 95% CI 1.07, 1.29), and “abnormal finding
163 on the examination of the lungs” (1.51, 95% CI 1.22, 1.87). These associations were further supported
164 by other MR methods (Figure S6). Additionally, we observed an association between a higher genetic
165 liability of OC and an increased risk of “ovarian cyst” (1.23, 95% CI 1.04, 1.46), as well as “lipoma
166 including lipoma of skin and subcutaneous tissues” (1.31, 95% CI 1.06, 1.61). Notably, the weighted
167 median and weighted mode MR methods specifically supported the association with “lipoma of the
168 skin and subcutaneous tissues” (Figure S6). There was some evidence supporting the association
169 between OC and higher risk of “postoperative infection” and “gram-negative septicemia”, identified
170 in the phenome-wide analyses on clear cell OC. However, these associations were observed only in
171 the IVW and MR-PRESSO MR methods (Figure S6). In the MR analyses on OC subtypes (Figure 2),
172 serous, endometrioid and clear cell OC were all associated with “secondary malignant neoplasm of
173 digestive systems” and “ascites”. Serous and endometrioid subtypes showed evidence of effects on
174 the respiratory outcomes. While the case numbers were lower for mucinous OC, this histotype was the
175 only one to show evidence of an association with “lipoma of the skin and subcutaneous tissues” as
176 evidenced by all MR methods. However, mucinous OC was not associated with any of the other
177 outcomes (Figure S6). The clear cell subtype was associated with “ovarian cyst”, “benign uterine
178 neoplasms”, “uterine leiomyoma”, and “disorders of menstruation”. Both serous and clear cell
179 subtypes increased the risk of postoperative infection traits, with clear cell also increasing risk for
180 “gram negative septicemia” (Figure 2 & Figure S6). We conducted analyses also excluding all known

181 OC cases, and associations with “secondary malignant neoplasm of digestive systems” and “ascites”
182 were attenuated, while estimates for all other outcomes were generally similar to the main findings
183 (Figure S7). The MR-Egger intercept did not suggest evidence of pleiotropy for most associations,
184 except for the association between serous OC and “postoperative infections” ($p_{\text{intercept}}=0.01$), which
185 therefore should be cautiously interpreted. No evidence of pleiotropy was found using the MR-
186 PRESSO outlier test and leave-one-out analysis (Figure S8). The estimates obtained from leave-
187 block-out analyses were broadly consistent with the overall findings (Figure S9).

188 MR analyses on serum biomarkers and physiological measures supported the association of
189 OC with lower FVC, endometrioid OC with lower oestradiol, clear cell OC with lower total
190 cholesterol, low density lipoprotein (LDL) cholesterol, and apolipoprotein B (ApoB), and mucinous
191 OC with higher gamma glutamyltransferase (GGT) and alanine aminotransferase (ALT) and lower
192 creatinine levels (Figure S10). Analyses excluding all known OC cases, provided generally similar
193 estimates to the main findings (Figure S11). For these identified associations, we did not find any
194 evidence of pleiotropy based on the MR-Egger intercept tests ($p_{\text{intercept}} \geq 0.13$) and leave-one-out
195 analyses. MR analyses on men provided evidence for a genetic association between mucinous OC and
196 ALT (Figure S10).

197 **3.2. Shared genetic architecture**

198 In the phenome-wide analyses assessing associations with individual OC risk-associated genetic
199 variant, five variants were associated with at least one of 15 (11 distinct) disease outcomes (all, $p \leq 4.5$
200 $\times 10^{-8}$, Figure 3, Table S6), and seven variants were associated with at least one of 22 biomarkers and
201 physiological measures (Figure S12). Some of these associations were also observed in the analyses
202 among men, with selected OC variants affecting conditions of the male genital organs (Figure 3,
203 Table S6 and Figure S12, Supplementary Results). To ensure the novelty of our findings, we
204 extensively checked the GWAS Catalog for any reports of associations with the identified (or related)
205 diseases involving the lead or proxy variants ($r^2 \geq 0.1$). As a result, we uncovered potential novel
206 associations between the OC risk-increasing allele from the *TERT* locus (rs10069690-T) and a lower

207 risk of “seborrheic keratosis” (0.79, 95% CI 0.72, 0.87), in females, with similar a finding observed in
208 males. To further validate the associations, we conducted colocalisation analysis at the *TERT* locus
209 (LD block of chr5:982252 to chr5:2132442, human genomic build of hg19) and identified rs4449583
210 as the candidate causal variant explaining around 70% of 0.9239 PPFC for OC-“seborrheic keratosis”
211 (Figure 4). This variant is also strongly associated with OC in the Phelan OC GWAS¹⁵ ($p = 7.8 \times 10^{-12}$),
212 and in LD with the index variant from this GWAS (rs7705526, $r^2 = 0.784$) and with the lead variant in
213 the primary discovery OC GWAS²² (rs100699690, $r^2 = 0.432$, Supplementary Methods). The
214 association between the *GPX6* locus and “disorders of iron metabolism” and “celiac disease” are
215 potentially novel, although colocalisation analysis did not suggest that OC share the same causal
216 variants with these two diseases (Figure S13). On the contrary, we identified independent potential
217 causal genetic variant for each “disorders of iron metabolism” (rs34409925, $r^2 = 0.207$ with
218 rs6456822) and “celiac disease” (rs3131101, $r^2 = 0.049$ with rs6456822) in further conditional analysis
219 (Figure S13).

220

221 4. Discussion

222 In this large-scale phenome-wide study, we aimed to explore comorbidity risks associated with OC,
223 with additional analyses to identify subtype-specific associations based on genetic variants reflecting
224 different OC histological types. We found evidence supporting a possible causal association between
225 OC and 12 disease outcomes, including “secondary malignant neoplasm of the digestive system”,
226 “ascites”, and respiratory disorders. We also identified some additional subtype-specific associations,
227 including clear cell OC showing links with “gram-negative septicemia”, “uterine leiomyoma” and
228 “disorders of menstruation”, and mucinous OC with “lipoma of the skin and subcutaneous tissue”.
229 Analyses using 43 biomarkers also identified an association with seven biomarkers and these findings,
230 along with those from disease analyses, may have important implications for improving the clinical
231 management of patients with OC. Moreover, our exploration of each OC-risk increasing variant in the
232 PheWAS revealed a set of traits that are genetically shared with OC. This analysis not only confirmed
233 previous GWAS findings but also shed light on novel associations deserving further investigations.

234 OC is the leading cause of cancer-related ascites in females.³³ Our MR-PheWAS study revealed
235 genetic evidence linking ascites to OC across all subtypes except mucinous OC, which is consistent
236 with the fact that over one-third of women with OC develop ascites during their lifetime.^{33, 34} Ascites
237 is commonly considered a sign of advanced OC,³⁵ and in line with this, the evidence for an
238 association in our analyses was removed after excluding known OC cases. In patients with OC, the
239 presence of ascites commonly indicates metastasis to peritoneal surfaces and reflects the extent of the
240 disease.³⁶ Our study further established a causal association between OC and “secondary malignant
241 neoplasms of the digestive system”, indicating intra-abdominal dissemination of the tumour. Again,
242 this association was sensitive to the inclusion of known OC cases in our analyses, and is likely to
243 explained by a metastasis of the OC affecting the gastrointestinal tract or, potentially, a new
244 synchronous gastrointestinal cancer.

245 OC has also been linked to respiratory complications.^{37, 38} Our study provides additional supporting
246 evidence for this association, specifically linking OC to an increased risk of “chronic airway

247 obstruction” and “abnormal lung examinations”. The findings were further supported by spirometry
248 measures, which showed a significant association between a higher genetic liability to OC and lower
249 FVC. While in women with OC lowering of lung function may follow from malignant pleural
250 effusion and tumour metastasis to the pleura³⁷ or complications related to surgical procedures,³⁸ in our
251 analysis the association with lung disease remained when OC cases were excluded, and the
252 spirometric association was seen both in women and in men, which may indicate a more profound
253 biological link.

254 OC as reflected by our genetic study arises from the epithelium of the ovaries, endometrium and
255 fallopian tubes, and while most of the lung also consists of epithelium, a direct link has not been well
256 established. Interestingly, colocalisation analyses in our study also suggested that disorder of skin
257 condition (seborrheic keratosis) shared a genetic locus (*TERT*, encoding the enzyme, telomerase) with
258 OC, highlighting potential shared pathomechanisms involving the epithelium, and implicating
259 telomerase function in the opposing risks of OC and the skin disorder.^{39,40}

260 Debulking cytoreductive surgery is the standard procedure for advanced OC. It involves the surgical
261 removal of female reproductive organs such as the ovaries, fallopian tubes, uterus, as well as
262 omentum and organs of the digestive system, such as the bowel, depending on the stage of OC.⁴¹ The
263 invasive nature of this intervention,⁴¹ combined with compromised immune system function resulting
264 from the disease,⁴² may explain the observed association between OC and “postoperative infection”
265 and “gram-negative septicemia”. It is important to note that these associations were predominantly
266 observed in clear cell OC, which has also been associated with “benign neoplasms of the uterus”,
267 “uterine leiomyoma”, and “disorders of menstruation”. The presence of these conditions, which
268 primarily involve the uterine tissue, could reflect the well-established association between
269 endometriosis and clear cell OC.^{43,44} Our study also identified OC subtype specific associations
270 between mucinous OC and “lipoma of the skin and subcutaneous tissue”. Regarding the biomarkers,
271 endometrioid OC was associated with lower serum oestradiol level, clear cell OC was associated with
272 lower serum total cholesterol, LDL cholesterol, and ApoB, while mucinous OC was associated with
273 higher serum GGT and ALT as well as lower creatinine levels. These findings suggests that

274 comorbidities and selected biomarkers may play a role in identifying and characterising different
275 subtypes of OC.

276 Beyond the causal link between OC and disease outcomes and biomarkers, genetic variants
277 contributing to a higher risk of OC may also be associated with other diseases and biomarkers. The
278 *ABO* locus is a well-known pleiotropic locus with established associations with cardiovascular and
279 cerebrovascular traits.⁴⁵⁻⁵⁰ Our study confirmed these existing findings, as we also observed
280 associations between the *ABO* variant and conditions such as “pulmonary heart disease”,
281 “thrombophlebitis”, and “other disorders of circulatory system”. Additionally, we confirmed the
282 association between *ABO* with many cardiometabolic biomarkers, including associations with markers
283 of glucose and liver metabolism, confirming the highly pleiotropic nature of this variant.⁵¹

284 Our study identified a significant association between the *GPX6* variant (rs6456822) and a lower risk
285 of disorders related to iron metabolism, which is a critical component of overall health.⁵² However,
286 while the lead variant rs6456822 showed this association, our colocalisation analysis identified an
287 independent potential causal variant, rs34409925, for “disorders of iron metabolism”. We also found
288 another potential causal variant at the *GPX6* locus, rs3131101, for “celiac disease”. The OC lead
289 variant rs6456822 was also associated with total protein levels in women, as well as with SHBG
290 levels in men. Our study also identified a potential novel association between the *TERT* locus
291 (rs10069690) and “seborrhic keratosis”. The follow up colocalisation analysis strengthened the
292 shared genetic structures between this skin condition and OC, further suggesting that rs444583 may
293 be a causal variant. Importantly, we confirmed previous associations of the *TERT* variant with uterine
294 leiomyoma^{44, 53} and total protein levels in women^{49, 51}, as well as with prostate cancer in men.⁵⁴ Our
295 findings highlight the value of using PheWAS for validating previous GWAS findings and identifying
296 potential novel variants for multiple traits including diseases and biomarkers that never been tested
297 before.⁵⁵

298 Our study has a number of strengths. Firstly, it is the largest (n =181,203) and most comprehensive
299 study to date that examines the health risks associated with a higher genetic susceptibility to OC.

300 Secondly, by including OC subtypes in our analysis, our study can identify subtype-specific
301 comorbidities, supporting more effective treatment and management. Lastly, our study extended the
302 phenome-wide analyses to each OC-risk-increasing variant. This approach enables us to confirm
303 previously identified variant-disease associations and also identify potential novel associations, thus
304 providing a more comprehensive understanding of the shared genetic links between OC and other
305 diseases. Further incorporation of colocalisation analysis is another strength of this study. It provided
306 evidence of a shared genetic locus (*TERT*) between OC and “seborrheic keratosis”, and identified a
307 potential novel causal variant, rs4449583, for this skin condition.

308 Our study also has some limitations. To ensure the validity of MR, the genetic instrument used must
309 be associated with the risk factor (in our case, OC), not associated with the confounders of the risk
310 factor-outcome association, and not influence the outcome except through its association with the risk
311 factor. While we used a genetic instrument that associates with OC and conducted our analysis on a
312 white-European population while adjusting for confounding factors, we cannot rule out the absence of
313 horizontal pleiotropic effects, which will affect interpretation and can bias the MR estimates.⁵⁶ To
314 address this concern, we repeated our analysis in men to further verify whether the association we
315 observed is genuinely related to OC or a reflection of genetic variants indicating shared mechanistic
316 pathways between both men and women. Another potential limitation of our study is that the UK
317 Biobank participants are relatively healthy compared to the general population,⁵⁷ introducing to
318 collider bias due to the “healthy volunteer” selection bias.⁵⁸ However, the modest collider bias in the
319 context of MR is less likely to affect MR results significantly when the selection into the study was
320 not strongly influenced by the risk factor.⁵⁸ Furthermore, the presence of biologically plausible causal
321 associations in our study provides support for the validity of our results despite this limitation. It could
322 be noted as a limitation of our study that our investigation into comorbidities associated with OC was
323 primarily conducted in a population predominantly not diagnosed with OC. However, sensitivity
324 analyses excluding all known OC cases confirmed that the associations with OC complications
325 relating to ascites and gastrointestinal metastases are related to the disease *per se*, while the other

326 observed disease and trait associations may reflect earlier stages of OC development or other
327 underlying biological differences.

328 In conclusion, our study identified robust associations between OC and comorbidities related to the
329 digestive and respiratory systems. Furthermore, we observed specific robust associations for different
330 subtypes of OC, such as “gram-negative septicemia”, “uterine leiomyoma” and “disorders of
331 menstruation” for clear cell OC, and lipoma for mucinous OC. Interestingly, our study points to a
332 shared genetic architecture between OC and a skin condition, leading to a discovery of a novel variant
333 which merits further investigation.

334

335 **Funding**

336 This study was supported by Medical Research Future Fund of Australia (MRFF), grant number
337 2007431.

338

339 **Competing interests**

340 The authors declare that they have no competing interests.

341

342 **Acknowledgements**

343 The authors are very grateful to the UK Biobank and UK Biobank Participants.

344

345 **Authors' contributions**

346 AM conceived the study with EH, conducted the analysis, and wrote the first draft. E.H. supervised
347 the study and advised on analyses. DS conducted the colocalisation analysis using summary results
348 generated by AM. All authors interpreted the results, revised the manuscript, and approved the final
349 manuscript.

350

351 **Ethical approval**

352 The study was conducted under UK Biobank application number 89630. UK Biobank study has
353 ethical approval from the National Information Governance Board for Health and Social Care and
354 Northwest Multicenter Research Ethics Committee (11/NW/0382).

355 **Consent to participate**

356 UK Biobank obtained informed consent from all participants to collect and use linked data for future
357 research use.

358 **Data availability**

359 This research has been conducted using the UK Biobank Resource under Application Number 89630.

360 The data and analytical code that support the findings of this study are available on application to the

361 UK Biobank (<https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>).

362

363 **References**

- 364 1. DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R,
365 Robbins AS, Jemal A. Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin*
366 2014;64:252-71.
- 367 2. Kurman RJ, Shih Ie M. The origin and pathogenesis of epithelial ovarian cancer: a
368 proposed unifying theory. *Am J Surg Pathol* 2010;34:433-43.
- 369 3. Sankaranarayanan R, Ferlay J. Worldwide burden of gynaecological cancer: the
370 size of the problem. *Best practice & research. Clinical obstetrics & gynaecology*
371 2006;20:207-25.
- 372 4. Perets R, Wyant GA, Muto KW, Bijron JG, Poole BB, Chin KT, Chen JY, Ohman
373 AW, Stepule CD, Kwak S, Karst AM, Hirsch MS, et al. Transformation of the fallopian tube
374 secretory epithelium leads to high-grade serous ovarian cancer in Brca;Tp53;Pten models.
375 *Cancer Cell* 2013;24:751-65.
- 376 5. Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. *The Lancet*
377 2019;393:1240-53.
- 378 6. Zhang S, Cheng C, Lin Z, Xiao L, Su X, Zheng L, Mu Y, Liao M, Ouyang R, Li
379 W, Ma J, Cai J, et al. The global burden and associated factors of ovarian cancer in 1990-
380 2019: findings from the Global Burden of Disease Study 2019. *BMC public health*
381 2022;22:1455.
- 382 7. Jiao YS, Gong TT, Wang YL, Wu QJ. Comorbidity and survival among women
383 with ovarian cancer: evidence from prospective studies. *Scientific reports* 2015;5:11720.
- 384 8. Lee L, Cheung WY, Atkinson E, Krzyzanowska MK. Impact of comorbidity on
385 chemotherapy use and outcomes in solid tumors: a systematic review. *Journal of clinical*
386 *oncology : official journal of the American Society of Clinical Oncology* 2011;29:106-17.
- 387 9. Bush WS, Oetjens MT, Crawford DC. Unravelling the human genome-phenome
388 relationship using phenome-wide association studies. *Nature Reviews Genetics* 2016;17:129-
389 45.
- 390 10. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation
391 studies: a guide, glossary, and checklist for clinicians. *BMJ* 2018;362:k601.
- 392 11. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott
393 P, Green J, Landray M, Liu B, Matthews P, et al. UK Biobank: an open access resource for
394 identifying the causes of a wide range of complex diseases of middle and old age. *PLOS*
395 *Medicine* 2015;12:e1001779.
- 396 12. Wei WQ, Bastarache LA, Carroll RJ, Marlo JE, Osterman TJ, Gamazon ER, Cox
397 NJ, Roden DM, Denny JC. Evaluating phecodes, clinical classification software, and ICD-9-
398 CM codes for phenome-wide association studies in the electronic health record. *PLoS One*
399 2017;12:e0175508.
- 400 13. King C, Mulugeta A, Nabi F, Walton R, Zhou A, Hyppönen E. Mendelian
401 randomization case-control PheWAS in UK Biobank shows evidence of causality for
402 smoking intensity in 28 distinct clinical conditions. *EClinicalMedicine* 2020;26:100488.
- 403 14. Verma A, Bradford Y, Dudek S, Lucas AM, Verma SS, Pendergrass SA, Ritchie
404 MD. A simulation study investigating power estimates in phenome-wide association studies.
405 *BMC bioinformatics* 2018;19:120.
- 406 15. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ,
407 Dennis J, Pirie A, Riggan MJ, Chornokur G, Earp MA, Lyra PC, Jr., et al. Identification of 12
408 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet*
409 2017;49:680-91.

- 410 16. Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, Anton-
411 Culver H, Chang-Claude J, Cramer DW, DiCioccio R, Dörk T, Goode EL, et al. A genome-
412 wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. *Nat*
413 *Genet* 2009;41:996-1000.
- 414 17. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K,
415 Widschwendter M, Vierkant RA, Larson MC, Kjaer SK, Birrer MJ, Berchuck A, et al. A
416 genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and
417 8q24. *Nat Genet* 2010;42:874-9.
- 418 18. Bolton KL, Tyrer J, Song H, Ramus SJ, Notaridou M, Jones C, Sher T, Gentry-
419 Maharaj A, Wozniak E, Tsai YY, Weidhaas J, Paik D, et al. Common variants at 19p13 are
420 associated with susceptibility to ovarian cancer. *Nat Genet* 2010;42:880-4.
- 421 19. Rafnar T, Gudbjartsson DF, Sulem P, Jonasdottir A, Sigurdsson A, Jonasdottir A,
422 Besenbacher S, Lundin P, Stacey SN, Gudmundsson J, Magnusson OT, le Roux L, et al.
423 Mutations in BRIP1 confer high risk of ovarian cancer. *Nat Genet* 2011;43:1104-7.
- 424 20. Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, Soucy
425 P, Fredericksen Z, Barrowdale D, Dennis J, Gaudet MM, Dicks E, et al. Genome-wide
426 association study in BRCA1 mutation carriers identifies novel loci associated with breast and
427 ovarian cancer risk. *PLoS genetics* 2013;9:e1003212.
- 428 21. Pharoah PD, Tsai YY, Ramus SJ, Phelan CM, Goode EL, Lawrenson K, Buckley
429 M, Fridley BL, Tyrer JP, Shen H, Weber R, Karevan R, et al. GWAS meta-analysis and
430 replication identifies three new susceptibility loci for ovarian cancer. *Nat Genet* 2013;45:362-
431 70, 70e1-2.
- 432 22. Kuchenbaecker KB, Ramus SJ, Tyrer J, Lee A, Shen HC, Beesley J, Lawrenson
433 K, McGuffog L, Healey S, Lee JM, Spindler TJ, Lin YG, et al. Identification of six new
434 susceptibility loci for invasive epithelial ovarian cancer. *Nat Genet* 2015;47:164-71.
- 435 23. Kelemen LE, Lawrenson K, Tyrer J, Li Q, Lee JM, Seo JH, Phelan CM, Beesley
436 J, Chen X, Spindler TJ, Aben KK, Anton-Culver H, et al. Genome-wide significant risk
437 associations for mucinous ovarian carcinoma. *Nat Genet* 2015;47:888-97.
- 438 24. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A,
439 Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, et al. The UK Biobank resource
440 with deep phenotyping and genomic data. *Nature* 2018;562:203-09.
- 441 25. Carroll RJ, Bastarache L, Denny JC. R PheWAS: data analysis and plotting tools
442 for phenome-wide association studies in the R environment. *Bioinformatics* 2014;30:2375-
443 76.
- 444 26. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan N, Thompson J. A
445 framework for the investigation of pleiotropy in two-sample summary data Mendelian
446 randomization. *Stat Med* 2017;36:1783-802.
- 447 27. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in
448 mendelian randomization with some invalid instruments using a weighted median estimator.
449 *Genet Epidemiol* 2016;40:304-14.
- 450 28. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data
451 Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*
452 2017;46:1985-98.
- 453 29. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal
454 pleiotropy in causal relationships inferred from Mendelian randomization between complex
455 traits and diseases. *Nat Genet* 2018;50:693-98.
- 456 30. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson
457 JR. Assessing the suitability of summary data for two-sample Mendelian randomization
458 analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol*
459 2016;45:1961-74.

460 31. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, Howson JMM. A
461 fast and efficient colocalization algorithm for identifying shared genetic risk factors across
462 multiple traits. *Nature communications* 2021;12:764.

463 32. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-
464 generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.

465 33. Parsons SL, Lang MW, Steele RJC. Malignant ascites: a 2-year review from a
466 teaching hospital. *European Journal of Surgical Oncology (EJSO)* 1996;22:237-39.

467 34. Ayantunde AA, Parsons SL. Pattern and prognostic factors in patients with
468 malignant ascites: a retrospective study. *Annals of Oncology* 2007;18:945-49.

469 35. Kipps E, Tan DS, Kaye SB. Meeting the challenge of ascites in ovarian cancer:
470 new avenues for therapy and research. *Nat Rev Cancer* 2013;13:273-82.

471 36. Ahmed N, Stenvers KL. Getting to know ovarian cancer ascites: opportunities for
472 targeted therapy-based translational research. *Front Oncol* 2013;3:256.

473 37. Porcel JM, Diaz JP, Chi DS. Clinical implications of pleural effusions in ovarian
474 cancer. *Respirology* 2012;17:1060-7.

475 38. Xu M, Zhang W, Gao C, Zhou Y, Xie Y. Postoperative pulmonary complications
476 and outcomes in cytoreductive surgery for ovarian cancer: a propensity-matched analysis.
477 *BMC anesthesiology* 2022;22:120.

478 39. Buckingham EM, Klingelhutz AJ. The role of telomeres in the ageing of human
479 skin. *Experimental Dermatology* 2011;20:297-302.

480 40. Chen B, Yan Y, Wang H, Xu J. Association between genetically determined
481 telomere length and health-related outcomes: A systematic review and meta-analysis of
482 Mendelian randomization studies. *Aging Cell* 2023;22:e13874.

483 41. Pomel C, Jeyarajah A, Oram D, Shepherd J, Milliken D, Dauplat J, Reynolds K.
484 Cytoreductive surgery in ovarian cancer. *Cancer Imaging* 2007;7:210-5.

485 42. Cubillos-Ruiz Juan R, Silberman Pedro C, Rutkowski Melanie R, Chopra S,
486 Perales-Puchalt A, Song M, Zhang S, Bettigole Sarah E, Gupta D, Holcomb K, Ellenson
487 Lora H, Caputo T, et al. ER Stress Sensor XBP1 Controls Anti-tumor Immunity by
488 Disrupting Dendritic Cell Homeostasis. *Cell* 2015;161:1527-38.

489 43. Mortlock S, Corona RI, Kho PF, Pharoah P, Seo JH, Freedman ML, Gayther SA,
490 Siedhoff MT, Rogers PAW, Leuchter R, Walsh CS, Cass I, et al. A multi-level investigation
491 of the genetic relationship between endometriosis and ovarian cancer histotypes. *Cell Rep*
492 *Med* 2022;3:100542.

493 44. Gallagher CS, Mäkinen N, Harris HR, Rahmioglu N, Uimari O, Cook JP, Shigesi
494 N, Ferreira T, Velez-Edwards DR, Edwards TL, Mortlock S, Ruhioğlu Z, et al. Genome-wide
495 association and epidemiological analyses reveal common genetic origins between uterine
496 leiomyomata and endometriosis. *Nature communications* 2019;10:4857.

497 45. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, Preuss
498 M, Stewart AF, Barbalic M, Gieger C, Absher D, Aherrahrou Z, et al. Large-scale association
499 analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet*
500 2011;43:333-8.

501 46. Lindström S, Brody JA, Turman C, Germain M, Bartz TM, Smith EN, Chen MH,
502 Puurunen M, Chasman D, Hassler J, Pankratz N, Basu S, et al. A large-scale exome array
503 analysis of venous thromboembolism. *Genet Epidemiol* 2019;43:449-57.

504 47. Klarin D, Busenkell E, Judy R, Lynch J, Levin M, Haessler J, Aragam K, Chaffin
505 M, Haas M, Lindström S, Assimes TL, Huang J, et al. Genome-wide association analysis of
506 venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular
507 disease. *Nat Genet* 2019;51:1574-79.

508 48. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, Hedman Å K,
509 Wilk JB, Morley MP, Chaffin MD, Helgadottir A, Verweij N, et al. Genome-wide

510 association and Mendelian randomisation analysis provide insights into the pathogenesis of
511 heart failure. *Nature communications* 2020;11:163.

512 49. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiha S, Narita A,
513 Konuma T, Yamamoto K, Akiyama M, Ishigaki K, Suzuki A, et al. A cross-population atlas
514 of genetic associations for 220 human phenotypes. *Nat Genet* 2021;53:1415-24.

515 50. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, Gagnon
516 DR, DuVall SL, Li J, Peloso GM, Chaffin M, Small AM, et al. Genetics of blood lipids
517 among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet*
518 2018;50:1514-23.

519 51. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars N, Benner C, Aguirre M,
520 Venkataraman GR, Wainberg M, Ollila HM, Kiiskinen T, Havulinna AS, Pirruccello JP, et
521 al. Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat Genet* 2021;53:185-
522 94.

523 52. Dev S, Babitt JL. Overview of iron metabolism in health and disease.
524 *Hemodialysis international. International Symposium on Home Hemodialysis* 2017;21 Suppl
525 1:S6-s20.

526 53. Rafnar T, Gunnarsson B, Stefansson OA, Sulem P, Ingason A, Frigge ML,
527 Stefansdottir L, Sigurdsson JK, Tragante V, Steinthorsdottir V, Styrkarsdottir U, Stacey SN,
528 et al. Variants associating with uterine leiomyoma highlight genetic background shared by
529 various cancers and hormone-related traits. *Nature communications* 2018;9:3636.

530 54. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ,
531 Dadaev T, Leongamornlert D, Anokian E, Cieza-Borrella C, Goh C, Brook MN, et al.
532 Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility
533 loci. *Nat Genet* 2018;50:928-36.

534 55. Pendergrass SA, Brown-Gentry K, Dudek SM, Torstenson ES, Ambite JL, Avery
535 CL, Buyske S, Cai C, Fesinmeyer MD, Haiman C, Heiss G, Hindorff LA, et al. The use of
536 phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype
537 relationships and pleiotropy discovery. *Genet Epidemiol* 2011;35:410-22.

538 56. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid
539 instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*
540 2015;44:512-25.

541 57. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, Collins R,
542 Allen NE. Comparison of sociodemographic and health-related characteristics of UK
543 Biobank participants with those of the general population. *Am J Epidemiol* 2017;186:1026-
544 34.

545 58. Gkatzionis A, Burgess S. Contextualizing selection bias in Mendelian
546 randomization: how bad is it likely to be? *Int J Epidemiol* 2019;48:691-701.

547

548

549 **Figure legends**

550 **Figure 1.** Manhattan plot showing the phenome-wide association of genetic risk scores of overall
551 (Panel A) and subtypes (Panel B) of ovarian cancer with wide ranges of diseases. Panel C summaries
552 the association of the top 15 signals identified at FDR threshold ($p = 0.0005$) in Panel A or Panel B
553 across the different genetic risk scores. The association for the first three diseases (in blue) is a
554 validation of the genetic risk scores. Green highlights indicate association that passed the FDR
555 threshold.

556 **Figure 2.** Mendelian randomisation causal estimates between ovarian cancer subtypes and selected
557 diseases outcomes. Estimates were from Inverse variance weighted methods. #SNPs are number of
558 genetic variants included in the analyses.

559 **Figure 3.** Phenome-wide association findings of ovarian cancer risk-increasing variants, with the
560 Manhattan plot showing selected variants associations across all the diseases outcomes (with the y-
561 axis showing the level of significance in $-\log_{10}$ scale), and the forest plot showing the selected variant
562 – disease associations estimate among women and men.

563 **Figure 4.** Colocalisation analysis of ovarian cancer with dermatological condition at the *TERT* locus.
564 Panel A. Stacked regional association plots of ovarian cancer and seborrheic keratosis. Panel B.
565 Portions of the Posterior probability for full colocalisation (PPFC) explained by each genetic variant
566 at the *TERT* locus. The r^2 colour legends reflect the r^2 of each variant with respect to the lead variant.

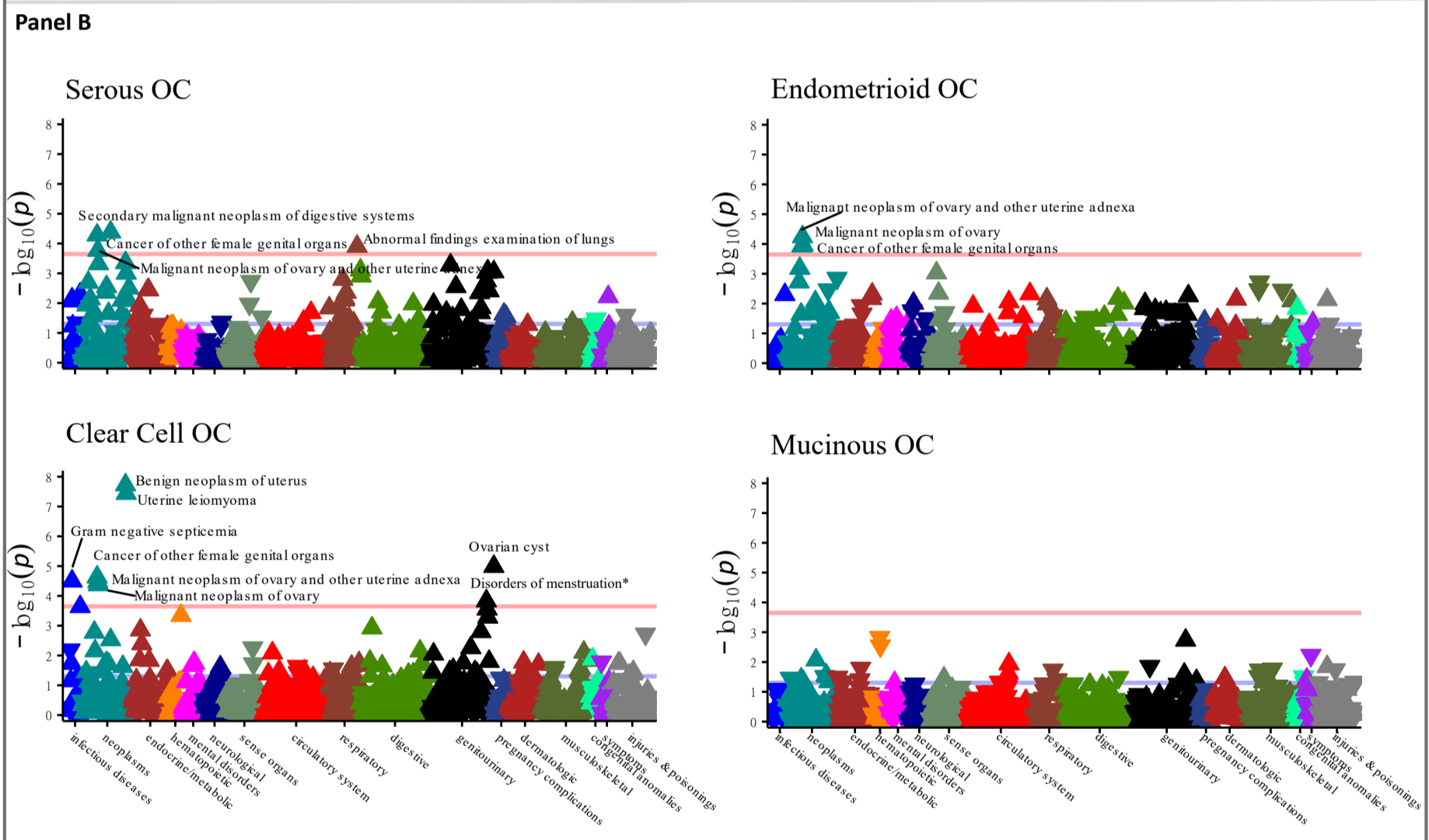
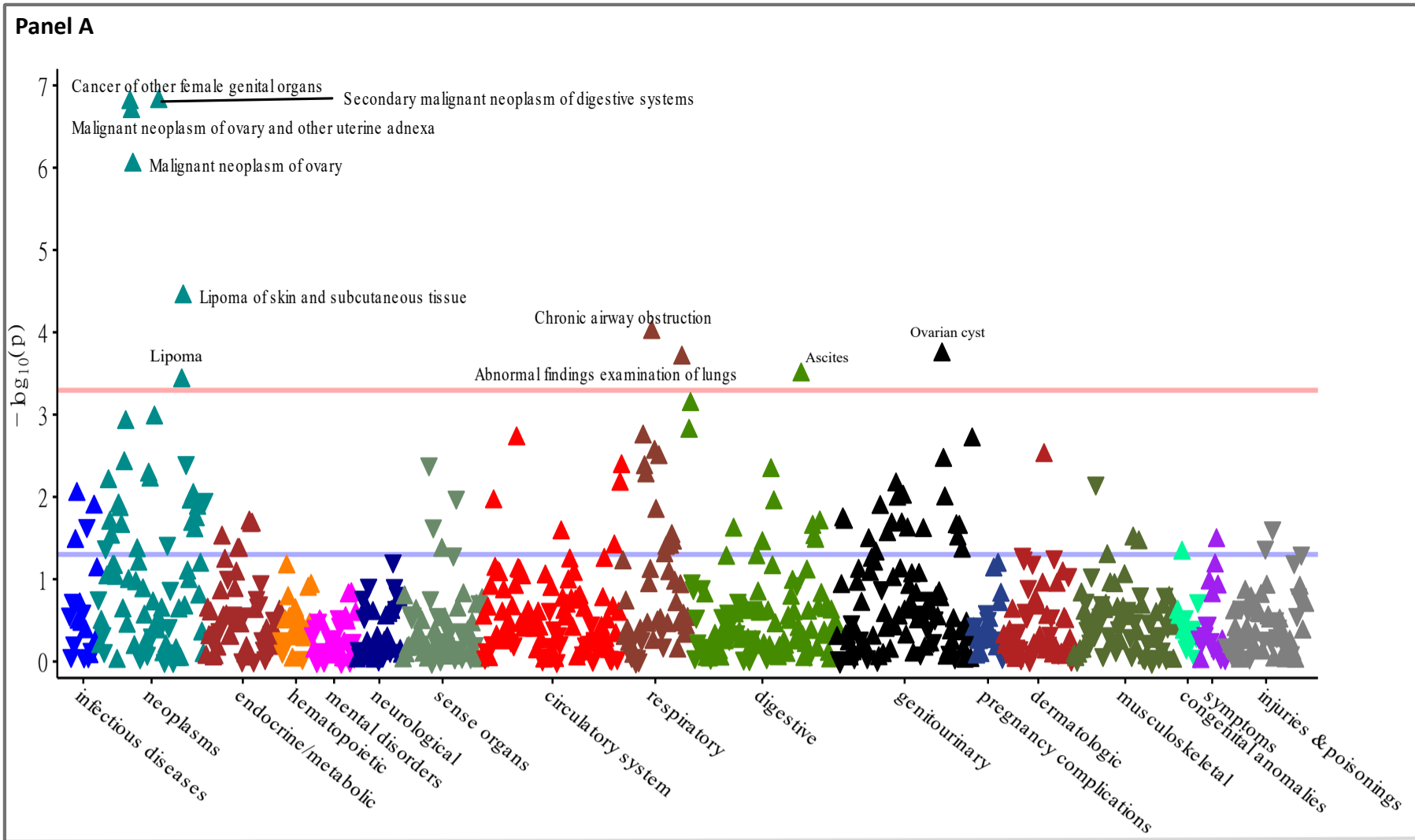
567

568 **Tables**569 **Table 1.** Prevalence of ovarian cancer by population characteristics

		All, n (%)	Ovarian cancer, n(%)	p
All		181,203	1,863 (1.3)	
Age (in years)	39–49	40,116 (22.2)	231 (0.6)	1.1 x 10 ⁻⁵¹
	50–59	62,041 (34.2)	556 (0.9)	
	60–73	79,046 (43.6)	1,076 (1.4)	
Education	None	31,059 (17.1)	421 (1.4)	0.06
	NVQ/CSE/A-levels	83,890 (46.3)	819 (1.0)	
	Degree/professional	64,785 (35.8)	607 (0.9)	
	Missing	1,469 (0.8)	16 (1.1)	
BMI (kg/m ²)	< 18.5	1,324 (0.7)	10 (0.8)	0.005
	18.5 - <25	70,988 (39.2)	643 (0.9)	
	25 - <30	66,596 (36.8)	713 (1.1)	
	≥ 30	41,746 (23.0)	489 (1.2)	
	Missing	549 (0.3)	9 (1.5)	
Alcohol consumption	Never	7,808 (4.3)	109 (1.4)	0.03
	Previous	6,526 (3.6)	64 (1.0)	
	Current	166,747 (92.0)	1,690 (1.0)	
	Missing	122 (0.1)	0 (0.0)	
General health	Excellent	31,630 (17.5)	286 (0.9)	1.0 × 10 ⁻¹⁶
	Good	109,012 (60.2)	1,041 (1.0)	
	Fair	33,519 (18.5)	405 (1.2)	
	Poor	6,431 (3.5)	112 (1.7)	
	Missing	611 (0.3)	19 (3.1)	

570 p is p-value generated from likelihood ratio test in logistic models adjusted for age, birth
571 location and assessment centers. BMI: body mass index.

572



Panel C

Disease outcome (Phecode)	Cases,n	Controls,n	OC				
			P	Serous OC P	Endometrioid OC P	Clear cell OC P	Mucinous OC P
Cancer of other female genital organs (184)	2388	122255	4.4E-07	5.2E-05	1.2E-04	2.3E-05	0.04
Malignant neoplasm of ovary and other uterine adnexa (184.1)	2091	122255	2.2E-07	1.7E-04	5.8E-05	3.4E-05	0.05
Malignant neoplasm of ovary (184.11)	2072	122255	1.0E-06	4.8E-04	1.1E-04	4.3E-05	0.04
Secondary malignant neoplasm of digestive systems (198.3)	1084	122255	2.0E-06	4.1E-05	0.009	0.003	0.69
Ascites (572)	845	104686	3.2E-04	0.01	0.007	0.04	0.66
Chronic airway obstruction (496)	5229	146930	1.0E-04	0.002	0.01	0.09	0.86
Abnormal findings examination of lungs (514)	1193	146930	6.9E-05	1.2E-04	0.12	0.07	0.68
Postoperative infection (80)	2153	164159	0.002	0.005	0.005	2.3E-04	0.78
Gram negative septicemia (38.1)	385	164159	0.007	0.007	0.52	3.2E-05	0.81
Benign neoplasm of uterus (218)	9234	122255	0.002	4.4E-04	0.60	1.9E-08	0.72
Uterine leiomyoma (218.1)	9008	122255	0.007	0.001	0.66	3.7E-08	0.93
Disorders of menstruation* (626)	15578	116934	0.03	0.002	0.22	1.5E-04	0.30
Ovarian cyst (628)	4147	116934	6.0E-05	8.9E-04	0.39	1.0E-05	0.92
Lipoma (214)	2636	122255	3.9E-04	0.02	0.04	0.12	0.03
Lipoma of skin and subcutaneous tissue (214.1)	2171	122255	3.7E-05	0.008	0.09	0.08	0.03

Secondary malignant neoplasm of digestive systems

Ovarian cancer (OC)	31	1.64 (1.33, 2.02)
Serous OC	20	1.34 (1.17, 1.55)
Endometrioid OC	5	1.82 (1.18, 2.81)
Clear Cell OC	5	1.55 (1.16, 2.06)
Mucinous OC	8	1.02 (0.82, 1.27)

Ascites

Ovarian cancer (OC)	31	1.48 (1.17, 1.86)
Serous OC	20	1.24 (1.07, 1.45)
Endometrioid OC	5	1.99 (1.06, 3.74)
Clear Cell OC	5	1.40 (1.01, 1.94)
Mucinous OC	8	1.04 (0.85, 1.26)

Chronic airway obstruction

Ovarian cancer (OC)	31	1.17 (1.07, 1.29)
Serous OC	20	1.10 (1.03, 1.17)
Endometrioid OC	5	1.29 (1.04, 1.59)
Clear Cell OC	5	1.12 (0.97, 1.29)
Mucinous OC	8	1.01 (0.94, 1.08)

Abnormal findings examination of lungs

Ovarian cancer (OC)	31	1.51 (1.22, 1.87)
Serous OC	20	1.28 (1.09, 1.51)
Endometrioid OC	5	1.43 (0.85, 2.43)
Clear Cell OC	5	1.29 (0.95, 1.74)
Mucinous OC	8	0.98 (0.84, 1.13)

Ovarian cyst

Ovarian cancer (OC)	31	1.23 (1.04, 1.46)
Serous OC	20	1.11 (0.99, 1.25)
Endometrioid OC	5	1.11 (0.79, 1.55)
Clear Cell OC	5	1.39 (1.08, 1.79)
Mucinous OC	8	0.99 (0.83, 1.18)

Lipoma

Ovarian cancer (OC)	31	1.24 (1.02, 1.52)
Serous OC	20	1.13 (0.98, 1.30)
Endometrioid OC	5	1.39 (0.89, 2.16)
Clear Cell OC	5	1.16 (0.90, 1.49)
Mucinous OC	8	1.14 (1.02, 1.27)

Lipoma of skin and subcutaneous tissue

Ovarian cancer (OC)	31	1.31 (1.06, 1.61)
Serous OC	20	1.16 (1.00, 1.34)
Endometrioid OC	5	1.34 (0.84, 2.16)
Clear Cell OC	5	1.20 (0.98, 1.47)
Mucinous OC	8	1.15 (1.03, 1.28)

Postoperative infection

Ovarian cancer (OC)	31	1.28 (1.10, 1.48)
Serous OC	20	1.15 (1.02, 1.29)
Endometrioid OC	5	1.61 (0.93, 2.78)
Clear Cell OC	5	1.47 (1.05, 2.05)
Mucinous OC	8	1.00 (0.89, 1.11)

Gram negative septicemia

Ovarian cancer (OC)	31	1.55 (1.06, 2.26)
Serous OC	20	1.29 (0.98, 1.69)
Endometrioid OC	5	1.26 (0.80, 2.62)
Clear Cell OC	5	2.74 (1.69, 4.45)
Mucinous OC	8	1.05 (0.80, 1.39)

Benign neoplasm of uterus

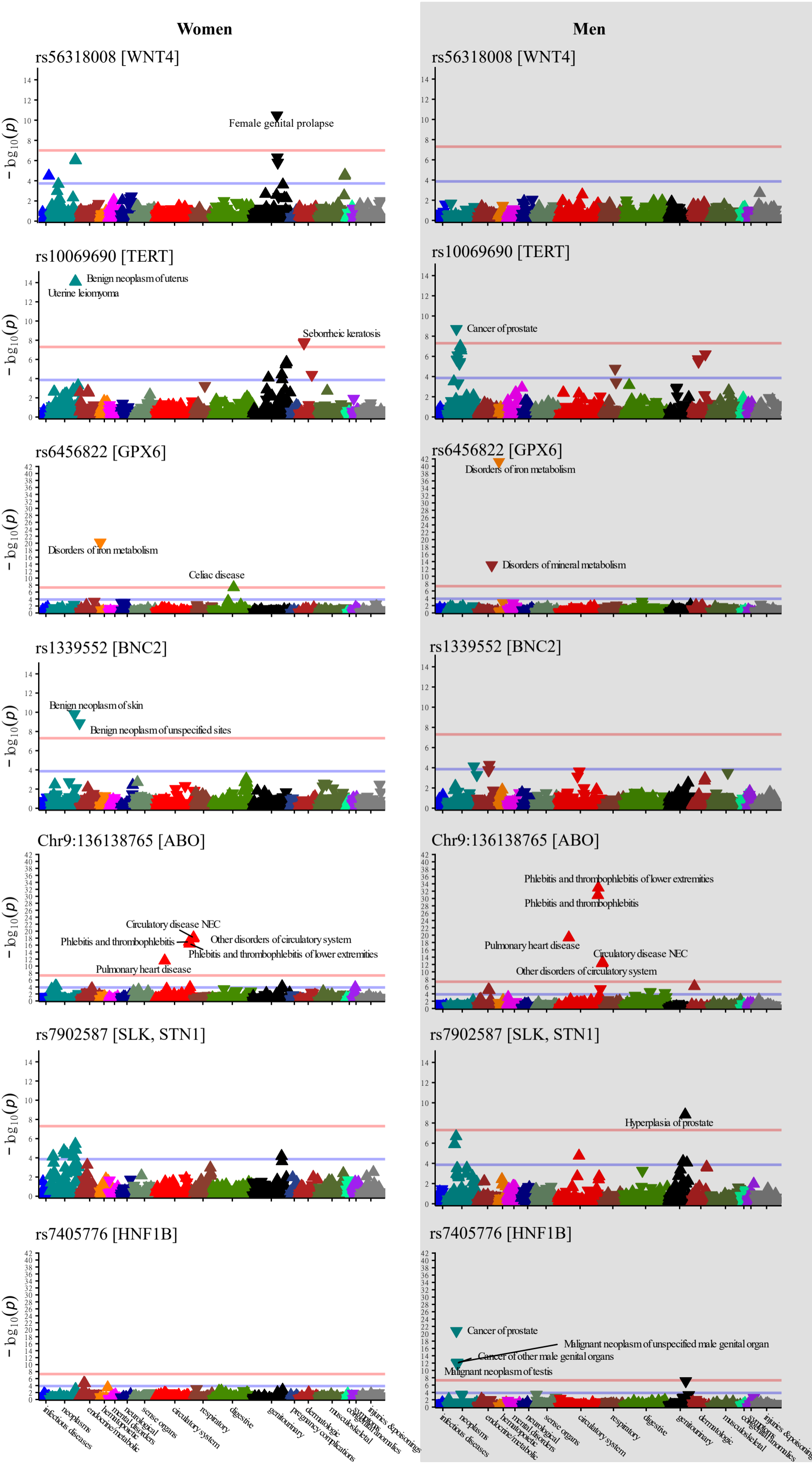
Ovarian cancer (OC)	31	1.13 (0.95, 1.33)
Serous OC	20	1.09 (0.96, 1.25)
Endometrioid OC	5	1.04 (0.66, 1.64)
Clear Cell OC	5	1.34 (1.09, 1.66)
Mucinous OC	8	1.01 (0.93, 1.10)

Uterine leiomyoma

Ovarian cancer (OC)	31	1.11 (0.94, 1.32)
Serous OC	20	1.09 (0.96, 1.24)
Endometrioid OC	5	1.03 (0.66, 1.63)
Clear Cell OC	5	1.34 (1.08, 1.66)
Mucinous OC	8	1.00 (0.92, 1.09)

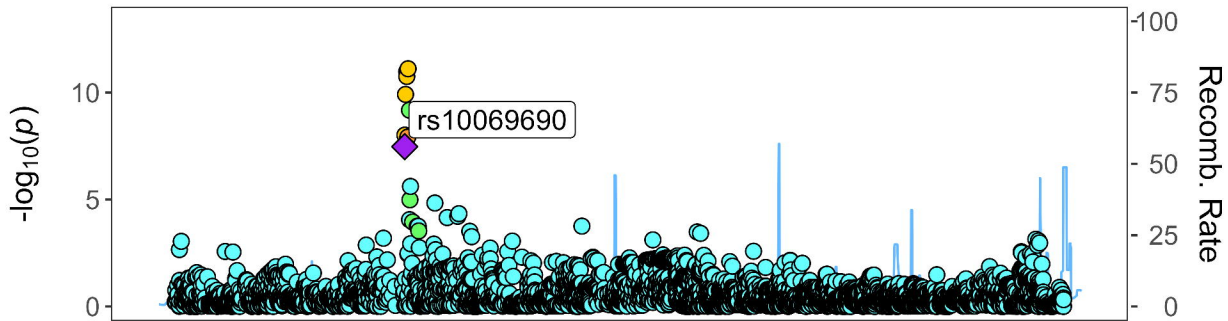
Disorders of menstruation

Ovarian cancer (OC)	31	1.06 (0.97, 1.16)
Serous OC	20	1.06 (0.99, 1.13)
Endometrioid OC	5	1.08 (0.91, 1.29)
Clear Cell OC	5	1.17 (1.05, 1.31)
Mucinous OC	8	1.02 (0.96, 1.09)

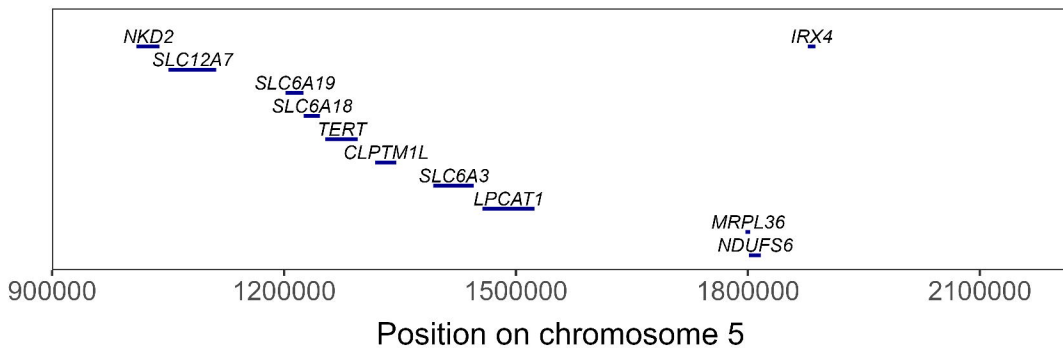
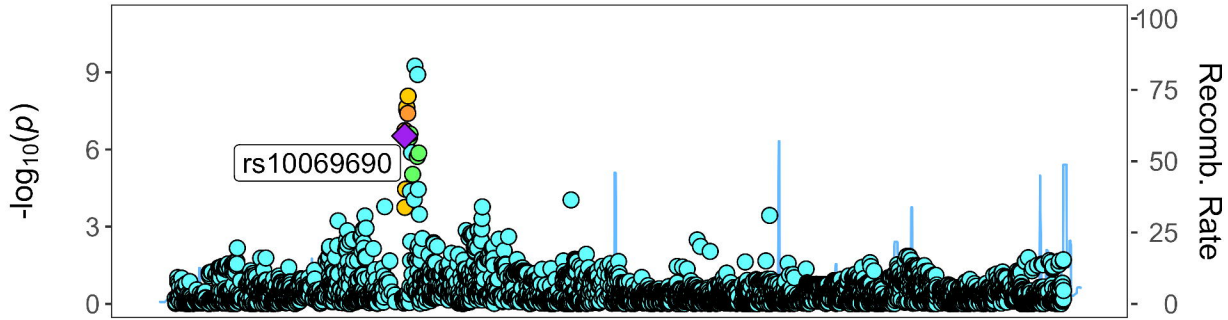


	Cases,n	Controls,n	OR (95% CI)
rs56318008 [WNT4]			
Female genital prolapse [618]	Women 10596	115919	0.86 (0.83, 0.90)
rs10069690 [TERT]			
Benign neoplasm of uterus [218]	Women 9234	122255	1.14 (1.10, 1.18)
Uterine leiomyoma [218.1]	Women 9008	122255	1.14 (1.11, 1.18)
Cancer of prostate [185]	Men 7073	105220	0.88 (0.85, 0.92)
Erythematous dermatosis [690]	Women 1603	156040	0.78 (0.72, 0.85)
	Men 1439	133564	0.81 (0.74, 0.88)
Seborrheic keratosis [702.2]	Women 1584	156040	0.78 (0.72, 0.85)
	Men 1389	133564	0.79 (0.72, 0.87)
rs6456822 [GPX6]			
Disorders of mineral metabolism [275]	Women 1110	143687	0.86 (0.79, 0.94)
	Men 1360	117777	0.74 (0.69, 0.80)
Disorders of iron metabolism [275.1]	Women 217	165527	0.41 (0.34, 0.49)
	Men 444	143918	0.40 (0.35, 0.45)
Celiac disease [557.1]	Women 1139	104682	1.30 (1.18, 1.42)
	Men 625	86330	1.10 (0.97, 1.24)
rs1339552 [BNC2]			
Benign neoplasm of unspecified sites [229]	Women 1397	119457	0.79 (0.73, 0.85)
	Men 674	109495	0.83 (0.74, 0.92)
Benign neoplasm of skin [216]	Women 5003	119457	0.88 (0.84, 0.91)
	Men 2969	109495	0.90 (0.85, 0.95)
Chr9:136138765 [ABO]			
Other disorders of circulatory system [459]	Women 7598	114822	1.20 (1.16, 1.25)
	Men 9446	86395	1.15 (1.11, 1.20)
Pulmonary heart disease [415]	Women 2199	114822	1.30 (1.21, 1.40)
	Men 2707	86395	1.35 (1.27, 1.45)
Circulatory disease NEC [459.9]	Women 7514	114822	1.21 (1.16, 1.26)
	Men 9320	86395	1.16 (1.11, 1.20)
Phlebitis and thrombophlebitis of lower extremities [451.2]	Women 1698	114822	1.42 (1.31, 1.54)
	Men 2196	86395	1.52 (1.41, 1.63)
Phlebitis and thrombophlebitis [451]	Women 1535	114822	1.44 (1.32, 1.56)
	Men 2056	86395	1.55 (1.45, 1.67)
rs7902587 [SLK, STN1]			
Hyperplasia of prostate [600]	Men 11270	112440	1.15 (1.10, 1.21)
rs7405776 [HNF1B]			
Cancer of prostate [185]	Men 7073	105220	0.84 (0.80, 0.87)
Cancer of other male genital organs [187]	Men 3331	111993	0.83 (0.78, 0.87)
Malignant neoplasm of unspecified* [187.1]	Men 3086	111993	0.82 (0.78, 0.86)
Malignant neoplasm of testis [187.2]	Men 3195	111993	0.82 (0.78, 0.87)

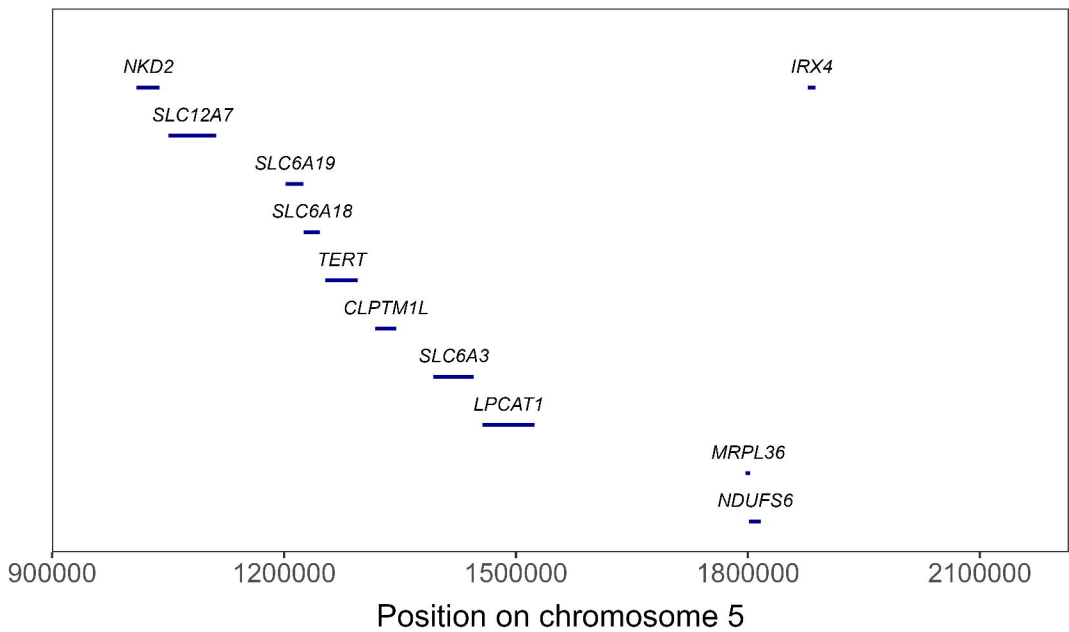
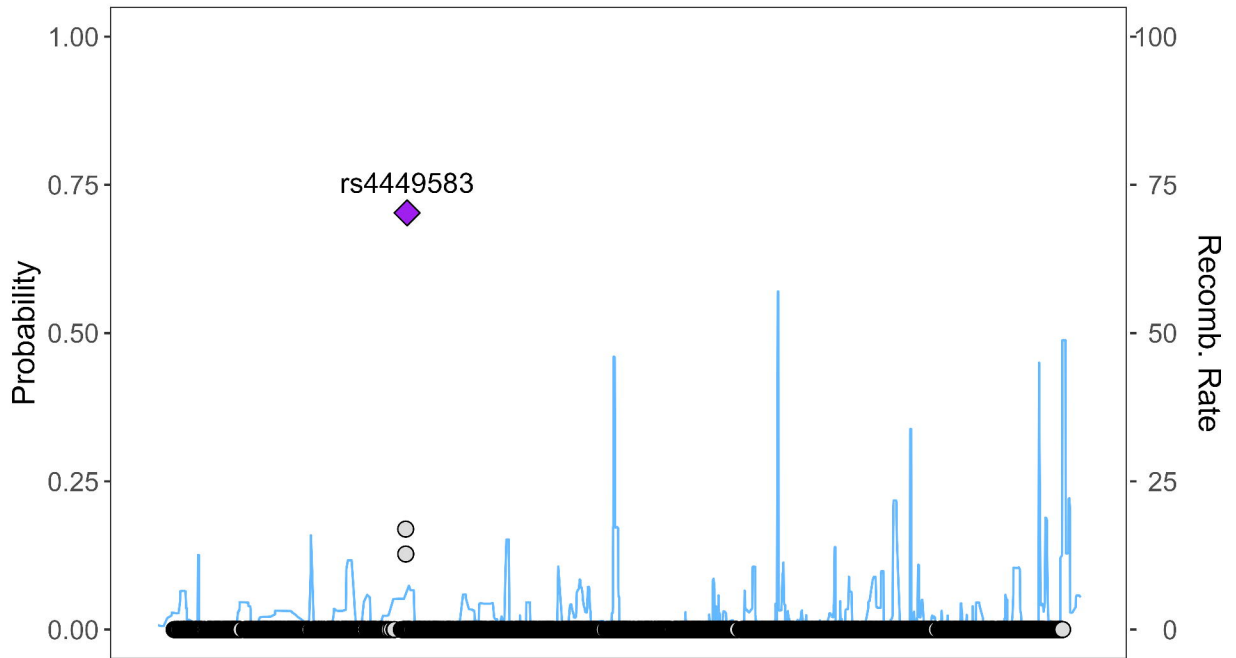
Ovarian cancer



Seborrheic keratosis



Ovarian cancer - seborrheic keratosis



Phenome-wide association study of ovarian cancer identifies common comorbidities and reveals shared genetics with complex diseases and biomarkers



181,203
female participants



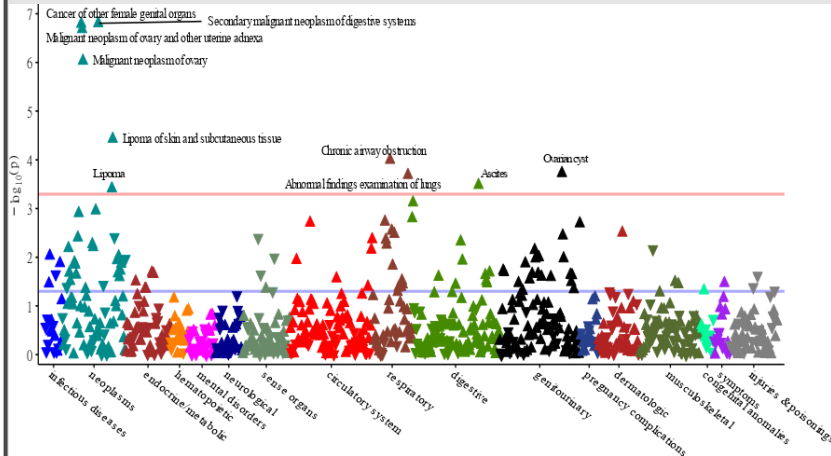
889
diseases
43
biomarkers and
physiological measures



35
ovarian cancer (OC) genetic variants
Genetic risk score (GRS)
for OC and the subtypes (serous, clear cell, endometrioid, and mucinous OC)

PheWAS validated the association between OC GRS and OC.

GRSs for OC or its subtypes were also associated with 12 other diseases, as well as 24 biomarkers and physiological measures



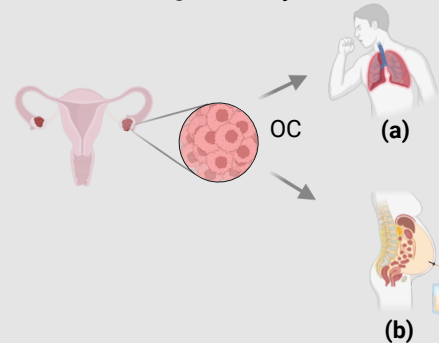
MR provided robust evidence of association between OC and the following diseases:

(a) Respiratory system

- Chronic airway obstruction
- Abnormal findings examination of the lungs
- Lower forced vital capacity (FVC)

(b) Digestive system

- Ascites
- Secondary malignant neoplasm of the digestive systems



PheWAS and Colocalisation analysis identified shared locus (*TERT*, rs4449583 potential causal variant) between OC and dermatological condition.

